

結合音樂與情緒之樂曲改編系統

洪宗貝
國立高雄大學
資訊工程系
教授
tphong@nuk.edu.tw

朱世祺
國立高雄大學
資訊工程系
學生
piliseric@gmail.com

邱國豪
國立高雄大學
資訊工程系
學生
roye1202456@yahoo.com.tw

簡嘉和
國立高雄大學
資訊工程系
學生
bcdef551920@hotmail.com

楊東祐
國立高雄大學
資訊工程系
學生
purits5832@hotmail.com.tw

摘要

現代科技蓬勃發展，機械化的製造，機械化的服務已經在我們生活中隨處可見，但是現代人的生活壓力也是水漲船高，卻沒有見到甚麼科技是直接對應人的心情，可以直接抒發或是與之產生共鳴，幫助人們排解心中壓力，音樂是最能直接反映人的心情，抒發人的感受的一項道具，所以我們想將人工智慧與音樂做結合幫助人們不只是外在更是包含內在的心靈。為了提供使用者一個可以符合使用者情緒的音樂軟體，本研究提出了一個以使用者情緒為改編基礎的樂曲改編系統，透過本系統讓音樂加入新的改動，使音樂不只侷限於原創的曲調。僅需要簡單的操作，便可將樂曲改編成符合使用者情緒的音樂，達到音樂舒壓之目的同時也達成音樂多元化的目標。

關鍵詞：音樂、情緒、改編

Abstract

This study furnishes a system, revising the music on the basis of the sentiment of the user, which modifies the music in order that it won't be restricted to the original tone.

With simple implementation, we can transform this music into another version fitting the emotion of the user for the purpose of stress relief and diversification of music.

Keywords: music, emotion, recompose

1. 前言

近年來隨著電腦的速度提升，網路的頻寬加大，人工智慧在各個範疇中都有很大的進展，讓電腦更加聰明並且具有學習能力，將會逐漸地成為在各個領域中嶄露頭角的一環，這將是勢必的趨勢。

音樂與技術結合的成品在現代來說越來越多，像是南科研究生發表的智慧型自動吉他伴奏系統之實現[7]將類神經網路搭配音樂，所寫出的智慧型吉他伴奏、淡江大學提出的應用調和搜尋演算法於結構最佳化之研究[1]利用調和搜尋演算法搭配音樂，將樂團演奏的音樂調和成最佳的樂章、利用基因演算法搭配編曲，撰寫出精彩的音樂，這些都已經是現有非常出色的研究。

本研究將朝這個方向前進，將這一項技術搭配上音樂撰寫的技巧，直接的影響在我們的生活跟我們息息相關的音樂上。讓他達到最貼近那個人的編寫風格，藉以達到讓人的心情感到舒適與抒發，可以對整個社會現代人的壓力得以紓解或是增進人們努力打拼的動力。

本研究主要分為四章節，第二章將探討有關音樂改編的文獻，第三章節將簡易介紹本系統的音符格式；第四章將會說明本系統之系統架構；最後，第五章將討論本研究的貢獻及未來方向。

2. 文獻探討

本章節將說明音樂改編相關的研究文獻。

2.1. 改編介紹

音樂術語中的改編，在 The New Grove Dictionary Of Music and Musicians[9]中將改編這個動作定義為 “*The reworking of a musical composition, usually for a different medium from that of the original.*” 意思相當於一個音樂的再創作，通常是用不同的表現方式展現。

2.2. 改編方式

在布梭尼《D 小調夏康舞曲》之樂曲分析與詮釋[5]中將布梭尼的改編手法概分為兩大類：第一類為原曲聲部織度與音域的更動，指的是在不會影響原曲結構為前提下，試圖將之改編以呈現更為豐富的音響效果之改編手法，其中包含八度音與和聲音的變化、速度的變化、擴大音域、使用先前樂段之素材、雙手齊奏、樂器音響的模仿以及、震音的使用等七項；第二類則是不同素材的添加，則是指將布梭尼額外添加與原曲不同之素材使其更具豐富性與炫技性的改編手法，包含了原有素材再發展及擴充、添加新的對位旋律、添加薩拉邦德舞曲節奏動機，以及添加炫技樂句等四項。

本系統採用音的變化以及速度改變對樂曲進行改編，其中音的變化可以加強樂曲的多樣性以及複雜度使音樂效果更加豐富，而速度的改變則會影響整首歌的所代表的情緒，像是激動的、生動的、平靜的……等，對音樂而言都具有明顯的改變。

3. 音符格式

本章節將說明本系統儲存音符的格式與相關運用。

3.1. 音符介紹

音樂可依照種種符號記載，使身處世界各地的人們可以演奏，亦可編寫創作，因此音樂才可以有跨地域性的交流。音符好比為音樂的文字，而樂譜便好比一首優美的詩詞。廣被熟知的樂譜變是由音符組成，經過了時代的變遷，

不斷的改良，才演變為現今普遍的五線譜架構。

音符在樂譜上所代表的意義在於其所對應樂器上的彈奏位置，而對我們而言，音符代表的是組成一首歌最小的一個單位也是本系統中最小組成單元，因此為了系統運用上的方便與不失其意義的情況下，經由歸納之後，我們將每一個音符以六個數字表示。

3.2. 音符特性與組成

在沈士育、鍾召鴻等學者所發表的鋼琴音樂之音符自動辨識[2]中提到音符的特性有音名與長短音，本系統以此兩個特性為基礎來加以延伸出可以代表一個完整音符的其他特性，經歸納後本系統使用六種不同的特性組成音符，分別為：音名、升降音、音域、連音、音長、附點。以下作詳細說明：

音名：古典樂理中將此類細分成音名與唱名，其中音名有七個，分別是 C、D、E、F、G、A、B，而唱名便是一般熟知的 do、re、mi、fa、so、la、si，由於彼此為對應關係，因此以音名來統稱。本系統中，以 0 代表休止符，1 代表 C、2 代表 D、3 代表 E、4 代表 F、5 代表 G、6 代表 A、7 代表 B。

升降：代表音符的音要比原來的音上升半個音或是下降半個音，以鋼琴來說明便是代表原來那個音所代表的鋼琴鍵的左邊黑鍵或是右邊黑鍵，右邊是升左邊是降。以 0 代表無升降，1 代表升，2 代表降。

音域：音域代表此音符所在的八度位置，以鋼琴而言，古典鋼琴一般會有八個八度，即是將鋼琴按鍵分為八個區塊，左邊區域的聲音聽起來較低右邊較高。本系統音域以 1 至 8 表示。

連音：連音代表與其他音符的關係，因此不能單獨存在，兩個音符之間若有連音記號，即被看過是一個音而不是分開的。系統中，0 代表沒有連音，1 代表此音符與下一個音符相連，2 代表此音符與上一個音符相連，3 代表此音符與上個音符和下一個音符相連。

音長：音長代表一個音符持續的時間長度。系統中以 0 代表十六分音符，1 代表八分音符，2 代表四分音符，3 代表二分音符，4 代表全音符。

附點：一個音符加了附點代表一個音符的音長比原來的多了一點五倍。系統中 0 代表無附點，1 代表有附點。

舉例來說一個在第五個八度四分音符的 do 即表示為 105020。

4. 樂曲改編系統

本系統流程分為兩大部分，第一部份為本系統中的輸入，亦為要進行下階段主要樂曲改編部份的前置作業，下個部分變是本系統的核心，分別有樂段切割、特徵點改編、爵士鼓伴奏與速度改編，此部份需先進行樂段切割部分，才可進行特徵點改編與速度改編，改編後可將音檔以 MIDI 音樂格式輸出。

各部分的詳細作法將介紹於下各小節。

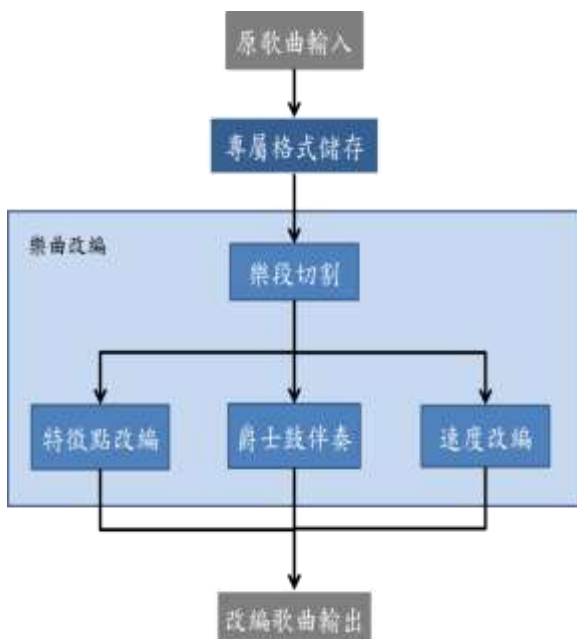


圖 1: 系統流程圖

4.1. 階段一：樂段切割

在王鴻文、劉志俊等學者提出的 MP3 音樂的聆賞情緒自動分類[1]中指出，由於一首音樂的長度太長，不容易分析，故將音樂按照小節為單位切割成許多主題片段，本系統依照此作法以小節為單位作改編後發現，改編的頻率過大，換言之，以小節為主切割出來的單位過小，且並非每小節都有其情緒特徵，導致改編後的樂曲結果不盡理想，因此本系統採取另一種樂段切割方法。

樂段切割方法考慮的是如何切取適當的樂曲長度，在作樂曲改編時，若不對其作切割動作，所有的運算將會以整首歌為單位，運算起來會十分複雜，且相對的運算量與運算時間

會較多，但若切割單位過小，則改編後的成效不彰，因此為了使樂曲進行後面特徵改編階段時簡便且具成效，本系統將一首完整的樂曲切割成多個樂段(Pattern)作為情緒特徵改編的基本單元，其中的 Pattern 代表的是歌曲中能夠帶有情緒意義的一段區間。

實際的作法分為三個步驟，以下詳細說明：

步驟一：利用每個段落會有一個「相對長音」，以此作為條件。作法是將所有音符的音長作一次統計，統計完後找到本次統計結果中音長最長的音符作為標的音符，以此音符作為切割條件。

步驟二：處理連音後的休止符部分，由於流行歌曲習慣以連音後加入休止符作為結尾，因此此休止符的地方也要將它認作需要切割的部分。

步驟三：根據步驟一與步驟二，將滿足切割條件的音符，該音符的下一個音，即為下一個 Pattern 的 Head。

圖 2 中，紅色的部分為滿足切割條件的音符的位置，藍色部分為該音符的下一個音，即為欲找尋的部分，紫色為其他音符。以此類推可以找到多個 Pattern 的 Head，在將 Head 的位置以陣列結構儲存。



圖 2：樂段切割示意圖

4.2. 階段二：情緒特徵改編

將分割好的樂曲片段進行旋律拍點的改編，這個部分本系統分別以兩個方法實作，一個是以一個小節長度為目標作的改編，另一個是以一拍為目標作的改編，藉由修改音樂旋律的拍點，以達到將樂曲旋律導向使用者所測訂的情緒方向，以下詳細說明：

作法一：以一個小節長度為目標作改編，根據高興與悲傷不同的情緒會有不同的樂曲特徵作修改。其中高興的特徵經由統計歸納後會有前兩拍音符數的音符總和會大於後兩拍的音符數總和的特性，而難過的特徵則有前兩拍音符的音符數總和會小於後兩拍音符數總和。以圖 3 而言，第一個小節前兩拍的音符數總和為四後兩拍則為二，第二個小節前兩

拍音符數總和為六後兩拍則為一，即符合高興的特徵。



圖 3：計算拍點總數示意圖

作法二：以一拍程度作為目標改編，根據高興、悲傷、冷靜、激昂不同情緒有不同的樂曲特徵作修改，將一拍分為四個位置，每個位置根據情緒給予不同的權重值，然後該一拍內的權重值總和需要達到該情緒一拍該有的門檻值。

4.3. 階段三：爵士鼓搭配

在此階段選擇加入爵士鼓作為伴奏元素。爵士鼓在現代音樂扮演著不可或缺的角色，由於鼓是打擊樂器，易於營造激動、急躁、和諧、詼諧等旋律樂器作不到的戲劇性效果，最能吸引觀眾注意，直接撼動人們的情緒，因此加入爵士鼓的伴奏對歌曲情緒加強與改變有顯而易見的成效。

在陳若涵學者所提出的以音樂內容為基礎的情緒分析與辨識[4]指出經過統計，一般歌曲在鼓的編曲上，較活潑輕快的歌曲前半部分便有鼓的編制且鼓點較為密集，而較平靜哀傷的歌曲後半部分才有鼓的編制且鼓點較為鬆散。本系統以此理論作為基準，並以爵士鼓中的大鼓、小鼓、Hit-hat 與 Crash 作為基礎單元，開發出爵士鼓生成演算法，此演算法一共分為五個步驟，以下作詳細說明：

步驟一、依據情緒要求生成基底的大鼓及 Hit-hat，在此步驟中將大鼓與 Hit-hat 依據前面的理論排好作為基底。

步驟二、隨機生成小鼓部份。

步驟三、調整步驟二的小鼓，避免大鼓和小鼓的聲音重疊。

步驟四、依照不同情緒要求調整 Crash 的位置。

步驟五、計算鼓點密度，會根據不同的情緒有不同的門檻值，通過門檻則完成，若沒通過便重回步驟一重新生成。

計算鼓點密度的公式如下：

$$DrumbeatDensity = \frac{BassdrumNum + SidedrumNum + HihatNum}{BarsNum}$$

其中，BassdrumNum：大鼓出現次數。

SidedrumNum：小鼓出現次數。

HihatNum：Hihat 出現次數。

BarsNum：小節總數。

公式 1：鼓點密度計算公式

4.4. 階段四：速度變換

在這階段我們藉由改變速度，將歌曲導向我們想要達到的情緒，因此調查現有歌譜的流行歌曲，並針對其情緒以速度加以分類，依照速度下去排序並統計出某一類情緒在不同速度上的歌曲數目分布以各速度的出現機率做為本系統改變的依據。

圖 4 圖 5 圖 6 圖 7 分別為四種情緒的速度分佈統計結果，橫軸為樂曲的速度單位為 bpm (beats per minute)，縱軸為介於該速度區間的歌曲數目，依據此統計結果，將音樂的速度依照欲改變的情緒作變化，舉例來說，若要將一首曲目往悲傷的方向作改變，其歌曲速度會以機率作改變，有五十三分之十七的機率速度會改變至五十到六十九之間，以此類推，可將歌曲依據歌曲情緒統計圖作不失一般性的情緒改編。

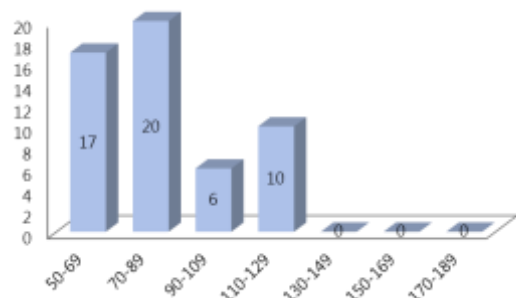


圖 4：悲傷歌曲速度分布圖

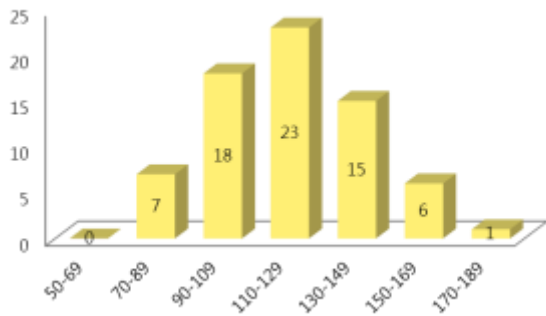


圖 5：高興歌曲速度分布圖

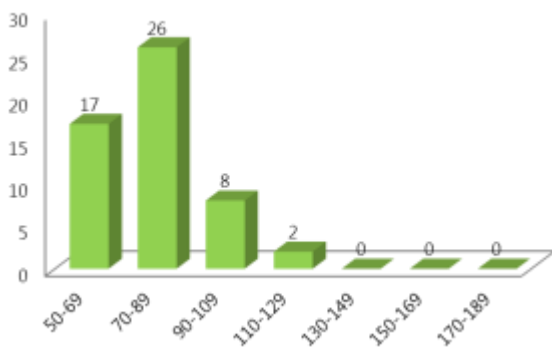


圖 6：平靜歌曲速度分布圖

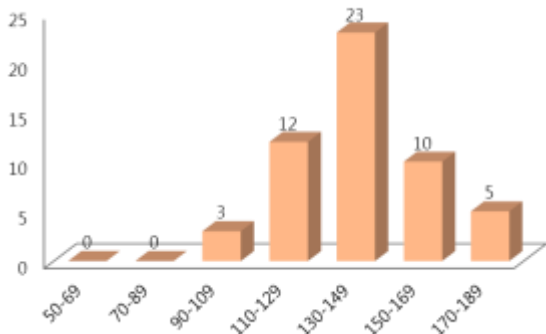


圖 7：激昂歌曲速度分布圖

5. 結論

本系統將整首音樂利用音符編碼的方式儲存在系統，進入樂曲改編階段後進行樂段切割，利用切割後的樂段作為音樂改編的輸入，輸入的樂段利用特徵點改編、爵士鼓伴奏及速度改編三種方法來達到音樂隨情緒改編之目的。

本系統可以多方應用到各個層面，只要選定好你現在的心情或是你想要的類型，不只是可以幫助人們排解無聊抒發心情，更可以在戲劇、電影、電視劇製作中節省了很多配樂成本，

最重要的是對於具有心理方面疾病的病人的音樂治療、抑或是對育懷孕中的母親所做的胎教以及未來嬰兒的音樂薰陶與智能啟發等等的方面，都可以藉由這一項系統，去對這方面做出應用達成一個真正回饋社會，為這個社會多出一份正面的幫助。

未來規劃用音樂辨識取代選擇既有的音樂檔作改編，換句話說，我們希望可以作到程式可以支援使用者可以直接輸入音樂網址或是輸入音樂檔，來進行情緒改編，如此自由度便會增加許多；目前還有想法是可以結合探測腦波的儀器，探測使用者是否因為在聆聽被我們改變心情的歌曲，心情而有所改變，這都是在未來能夠發展的方向。

參考文獻

- [1]王鴻文、劉志俊 2010, “MP3 音樂的聆賞情緒自動分類”, 中華大學資訊工程系
- [2]沈士育、鍾召鴻 2002, “鋼琴音樂之音符自動辨識”
- [3]李重光 2008, 音樂理論基礎——最權威的中文樂理教本
- [4]陳若涵 2006, “以音樂內容為基礎的情緒分析與辨識”, 國立清華大學資訊系統與應用所
- [5]陳晏萱、歐玲如 2014, “布梭尼《D 小調夏康舞曲》之樂曲分析與詮釋 An Analysis and Interpretation of F. Busoni-J.S. Bach's *Chaconne in D Minor*”
- [6]張永康、陳聖中 2013, “應用調和搜尋演算法於結構最佳化之研究”
- [7]趙樹堂、黃建生 2008, “智慧型自動吉他伴奏系統之實現”
- [8]Chris Dobrian 1993, Music and Artificial Intelligence
- [9]Grove, George 1878, The New Grove Dictionary Of Music and Musicians
- [10]Juslin, Patrik N./ Sloboda, John 2011 ,Handbook of Music and Emotion: Theory, Research, Applications
- [11] K.T.Strongman 1993, 情緒心理學
- [12]Stuart Russell, Peter Norvig 2002, Artificial Intelligence: A Modern Approach(2nd Edition),
- [13]Tao Li, Mitsunori Ogihara 2003 “Detecting Emotion in Music”, Department of Computer Science

[14]Yi-Hsuan Yang, Homer H. Chen, 2011 ,
“*music emotion recognition*”,CRC press